THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 60, No. 9, November 1981
Printed in U.S.A.

A Unified Theory of Data-Aided Equalization

By M. S. MUELLER and J. SALZ

(Manuscript received March 17, 1981)

A unified theory is presented for data-aided equalization of digital data signals passed through noisy linear dispersive channels. The theory assumes that some past and/or future transmitted data symbols are perfectly detected. We use this hypothesis to derive the minimum mean-square error receiver. The optimum structure consists of a matched filter in cascade with a transversal filter combined with a linear intersymbol interference canceler which uses the ideally detected data symbols. The main result is an expression for the optimized mean-square error as a function of the number and location of the canceler coefficients, the s/n, and the channel transfer function. When the number of canceler coefficients is zero, we get the well-known result for linear equalization. When the causal or postcursor canceler approaches infinite length, we obtain the well-known decision feedback result. When both the precursor and postcursor cancelers become infinite, we obtain the very best result possible, namely, the matched-filter bound dictated from fundamental theoretical considerations. Neither the decision feedback nor the matched-filter results can be achieved in practice since their implementation requires infinite memory and storage. Our theory can be used to calculate the rate of approach to these ideals with finite cancelers.

I. INTRODUCTION

The theory of linear and decision feedback equalization to mitigate the effects of intersymbol interference (ISI) and noise in digital data transmission is well known.[1-4] In this paper, the problem of equalization is cast in a general framework of an ISI canceler aided by past and/or future data values. This general structure is suggested from optimal detection theory and is shown in Fig. 1. The optimal detector of digital data in the presence of additive Gaussian noise and ISI is comprised of a matched filter and an ISI estimator which is used to cancel the

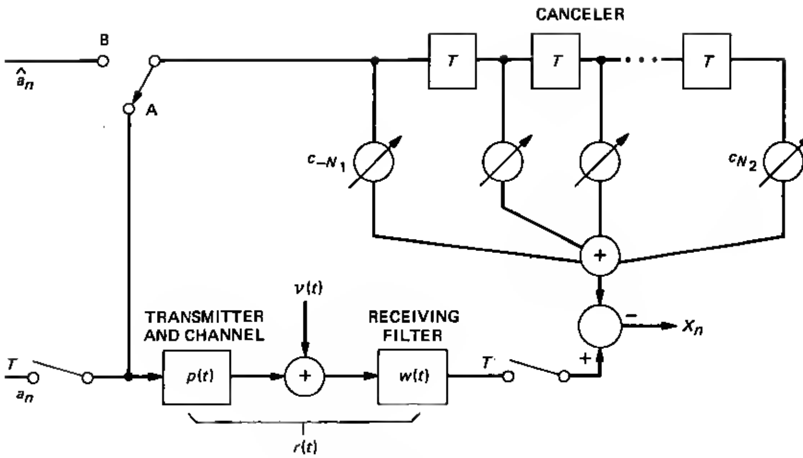

Fig. 1—Block diagram of data-aided equalization.

interference.[5,6] The implementation of this structure is often impractical because of its complexity.[7,8]

In our theory we postulate that some portion of the ISI can be perfectly synthesized and, therefore, subtracted from the incoming signal. In other words, we replace the optimal estimator with a practical one. The effect of the remaining interference is then minimized by a linear filter or a conventional linear equalizer. In practical systems, however, perfect estimation cannot be achieved; therefore, our results serve as ideal limits. The inclusion of occasional errors in our theory has proved mathematically intractable so far.

In Section II, we determine the minimal mean-square error (mse) when an arbitrary set of data symbols is known to the receiver. In Section III, the optimal receiving filter is derived and analyzed. The performance of the infinite linear equalizer, the decision feedback equalizer and the infinite canceler are obtained as special cases of the general result. Section IV covers a discussion of numerical results.

II. MINIMUM MSE FOR DATA-AIDED EQUALIZATION

In Fig. 1, the transmitter generates the data sequence $\{a_n\}$ whose elements are assumed to be independent identically distributed (i.i.d.) discrete random variables. These discrete amplitudes sequentially modulate the pulse $p(t)$ at a rate $1/T$ to produce the transmitted signal. The pulse shape, $p(t)$, can be viewed as the overall impulse response of the transmitting filter and the transmission channel. White noise, $\nu(t)$, is added to the received signal which is then applied to the linear receiving filter, $w(t)$. The output signal is sampled at the symbol rate $1/T$ and combined with the output of the canceler. The linear

canceler is modeled as a transversal filter with coefficients, $\{c_n\}$, where $n \in S$, and where the set of integers S denotes the range of the canceler's taps.

For most applications, and for the special cases investigated later in this Section, the range will contain tbe neighboring taps of the reference location but never the reference location itself, and consequently, $S = \{-N_1, \cdots, -1, 1, \cdots, N_2\}$. The canceler operates on the past received symbols $a_{n-1}, \cdots, a_{n-N_2}$ and on the future received symbols $a_{n+1}, \cdots, a_{n+N_1}$, which are assumed to be known to the receiver. Clearly, to realize an operation on the future data symbols, a time delay of at least N_1T seconds has to be introduced.

For a general set S the output signal $x_n = x(nT)$ can, thus, be written as

$$x_n = \sum_{k=-\infty}^{\infty} r_k a_{n-k} - \sum_{k \in S} c_k a_{n-k} + \xi_n . \tag{1}$$

Where $r_k = r(kT)$ is the overall impulse response evaluated at $t = kT$,

$$r_k = T \int_{-\infty}^{\infty} w(\tau) p(kT - \tau) d\tau , \tag{2}$$

and where $\xi_k = \xi(kT)$, i.e.

$$\xi_k = \int_{-\infty}^{\infty} w(\tau) \nu(kT - \tau) d\tau, \quad \text{for} \quad k = -\infty, \cdots, -1, 0, 1, \cdots, \infty. \tag{3}$$

To facilitate modeling of various types of linear modulation scbemes, the data sequence, the noise, and all impulse responses are assumed to be complex valued. In general, $p(t)$ will be the preenvelope of the passband transmission system with respect to a carrier frequency. This notation has become extremely useful and economical in this field.[9] Specifically, it permits a unified presentation of baseband and passband systems.

Tbe output signal, x_n, after slicing or quantizing is usually taken to be an estimate of the transmitted data symbol a_n. Our goal now is to obtain a receiving filter, $w(t)$, and canceler taps, $\{c_n\}$, so that the mse,

$$\epsilon = E\{|x_n - a_n|^2\} , \tag{4}$$

is a minimum. To determine the optimal canceller coefficients, $\{c_n\}$, we differentiate eq. (4) with respect to c_n, $n \in S$, and set the result to zero

$$\frac{\partial \epsilon}{\partial c_n} = \sigma_a^2 [c_n^* - r_n^* + c_n - r_n] = 0, \quad \text{for} \quad n \in S , \tag{5}$$

where

$$E\{a_n a_k^*\} = \sigma_a^2 \delta_{n,k}\,, \tag{6}$$

and $\delta_{n,k}$ is the Kronecker delta. The immediate conclusion from eq. (5) is that for $n \in S$

$$c_n = r_n\,. \tag{7}$$

Inserting this into eq. (1), we get

$$x_n = \sum_{k \notin S} r_k a_{n-k} + \int_{-\infty}^{\infty} w(\tau)\nu(nT - \tau)d\tau\,. \tag{8}$$

Thus far, our approach is perfectly obvious. By knowing the data symbols for all integers $k \in S$, it is possible to synthesize the resulting ISI associated with these symbols and subtract it from the current signal sample x_n. If the set S contains all integers $k < n$, we use all the already-decided-upon data symbols (available at the receiver without delay) to synthesize the postcursor ISI. This is precisely what is done in decision feedback equalization. If the set S contains all the integers, except the one associated with the present instant, n, all ISI is eliminated. But this, of course, requires infinite delay. In practice, the set S will be finite and our main concern will be to determine bow it influences the mse.

We now proceed to optimize the receiving filter, $w(t)$, for a given set S. Inserting eq. (8) into eq. (4) and using eq. (6), the resulting mse can be expressed as

$$\epsilon = \sigma_a^2 \left[\sum_{k \notin S} |r_k|^2 - r_0 - r_0^* + 1 + \sigma^2 \right], \tag{9}$$

wbere

$$E\{\nu(t)\nu(t+\tau)^*\} = \sigma_\nu^2 T\delta(\tau)\,, \tag{10}$$

and where

$$\sigma^2 = T\frac{\sigma_\nu^2}{\sigma_a^2}\int_{-\infty}^{\infty} |w(t)|^2 dt\,. \tag{11}$$

We remark that more general noise covariances can be included, but the calculations become more cumbersome without yielding additional insights.

To obtain the optimum $w(t)$, let

$$w(t) = w_0(t) + \lambda\mu(t) \tag{12}$$

and define

$$U_k = T\int_{-\infty}^{\infty} w_0(\tau)p(kT - \tau)d\tau\,, \tag{13}$$

where $w_0(t)$ is the optimum impulse response of the receiving filter and where the U_k are the samples of the optimized overall impulse response. It follows immediately from eq. (2) that

$$r_k = U_k + \lambda T \int_{-\infty}^{\infty} \mu(\tau) p(kT - \tau) d\tau . \tag{14}$$

When

$$\left. \frac{\partial \epsilon}{\partial \lambda} \right|_{\lambda=0} = 0 \tag{15}$$

is calculated from eq. (9), we obtain an equation for the optimum receiving filter:

$$w_0(t) N_0 = p(-t)^* - \sum_{k \not\in S} U_k p(kT - t)^*, \tag{16}$$

where

$$N_0 = \frac{\sigma_\nu^2}{\sigma_a^2} . \tag{17}$$

The interpretation of eq. (16) is standard: the optimum receiving filter is comprised of a matched filter $p(-t)^*$ in cascade with a transversal filter having taps only at those locations where the canceler has none. This structure is shown in Fig. 2.

To obtain our central result, an expression for the optimal mse, we multiply eq. (16) by $Tw_0(t)^*$ and integrate from $-\infty$ to $+\infty$. This yields with the aid of eqs. (9), (11), and (13) the result

$$\epsilon_{\text{opt}} = \sigma_a^2 (1 - U_0) . \tag{18}$$

The explicit determination of U_0 is the subject of the next section.

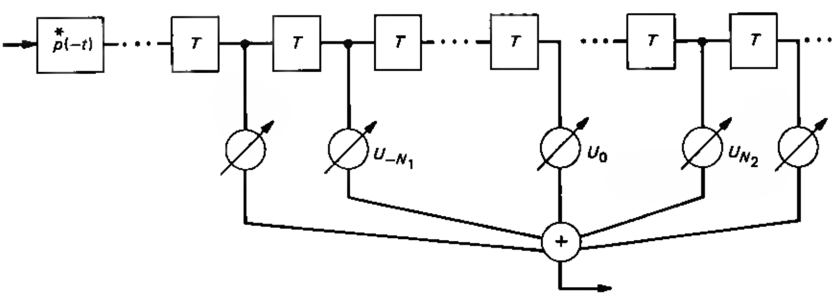

Fig. 2—Optimal receiving filter, $w_0(t)$.

III. THE OPTIMAL FILTER

To determine the sample values, $\{U_m\}$, of the optimal overall impulse response, define the autocorrelation function of the channel impulse response as

$$R_m = T \int_{-\infty}^{\infty} p(mT - \tau) p(-\tau)^* d\tau \tag{19}$$

and its Fourier transform as

$$R(\omega) = \sum_{m=-\infty}^{\infty} R_m e^{-jm\omega T}. \tag{20}$$

After multiplying eq. (16) by $Tp(nT - \tau)$ and integrating from $-\infty$ to $+\infty$, we get the following system of linear equations in $\{U_m\}$,

$$U_m N_0 = R_m - \sum_{k \notin S} U_k R_{m-k} \quad \text{for} \quad \text{all } m. \tag{21}$$

To determine the optimal receiving filter, we only need to know U_m for $m \notin S$. From eq. (21) we extract the equations necessary to determine U_0 and partition them as follows

$$U_0(N_0 + R_0) = R_0 - \sum_{k \notin J} U_k R_{-k} \quad \text{for} \quad m = 0, \tag{22}$$

$$\sum_{k \notin J} U_k M_{m-k} = (1 - U_0) R_m \quad \text{for} \quad m \notin J, \tag{23}$$

where we defined

$$M_k = R_k + N_0 \delta_{k,0}, \tag{24}$$

and where J is the set S augmented by $m = 0$.

Note that the indices of the unknowns and the indices of the right-hand sides of eq. (23) have gaps of the same size and at the same locations. (See the definition of J above.) Thus, the set of equations in (23) is not in a standard form and the solution technique is not obvious. In Appendix A we develop a technique to solve this infinite set of equations witb finite gaps. It involves the solution of a special infinite set of equations without gaps. To compensate for the gaps, we augment the original set of equations. Specifically, we add for $m \in J$ a finite number of equations to the infinite set such that the solution vanishes for $m \in J$. From Appendix A, we determine that the optimum mse becomes

$$\epsilon_{\text{opt}} = \sigma_a^2 \frac{N_0}{N_0 + H_0}, \tag{25}$$

where H_0 is determined from the following set of equations:

$$\sum_{m \in J} M_{k-m}^{-1} H_m = \delta_{k,0} - N_0 M_k^{-}, \quad \text{for} \quad k \in J \tag{26}$$

and where M_k^{-1} is the inverted sequence of M_k, i.e., it satisfies

$$\sum_n M_n M_{k-n}^{-1} = \delta_{k,0} \quad \text{for} \quad \text{all } k. \tag{27}$$

In the following section, we investigate the minimal mse for some special cases. As mentioned initially, we use the realistic assumption that the set S contains the neighboring locations $\{-N_1, \cdots, -1, 1, \cdots, N_2\}$. Then $J = \{-N_1, \cdots, N_2\}$ and the coefficient matrix in eq. (26) is a finite Toeplitz matrix. The solution of eq. (26) and, thus, H_0 is unique and is guaranteed to exist when $R(\omega) + N_0$ is bounded away from zero and infinity.[10] These conditions are very mild and are satisfied in most cases of practical interest.

3.1 Infinite length equalizer

For $N_1 = N_2 = 0$ the set J includes only the zero integer and all canceler coefficients vanish. Consequently, eq. (26) degenerates to a single equation

$$H_0 M_0^{-1} = 1 - N_0 M_0^{-1}. \tag{28}$$

Solving this for H_0 and inserting it into eq. (25), we obtain the standard result for the optimum linear equalizer,[2]

$$\epsilon_{\text{opt}} = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} \frac{N_0 d\omega}{R(\omega) + N_0}, \tag{29}$$

where we expressed M_0^{-1} in terms of its Fourier transform

$$M_0^{-1} = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} \frac{d\omega}{R(\omega) + N_0}. \tag{30}$$

3.2 Matched-filter

When $N_1 = N_2 = \infty$ the set J is infinite; i.e., the canceler subtracts all the ISI. Equation (26) in this case yields

$$\sum_{m=-\infty}^{\infty} M_{k-m}^{-1}(H_m + N_0 \delta_{m,0}) = \delta_{k,0} \quad \text{for} \quad \text{all } k. \tag{31}$$

Comparing this with eq. (27) gives the result

$$M_m = H_m + N_0 \delta_{m,0}. \tag{32}$$

From eqs. (32) and (23) it follows that $H_m = R_m$ for all m. Thus, the

optimum mse for this case is

$$\epsilon_{\text{opt}} = \sigma_a^2 \frac{N_0}{N_0 + R_0}. \tag{33}$$

This is recognized as the matched-filter bound for the optimal detection of a known signal in noise and in the absence of ISI.

3.3 One-sided canceler of infinite length

For $N_1 = 0$ and $N_2 = \infty$ tbe canceler performs as an ideal decision feedback equalizer of infinite length. In Appendix B the mse is derived for tbe more general case $N_1 \neq 0$, i.e. for a decision feedback equalizer with a limited number of noncausal taps. The result is

$$\epsilon_{\text{opt}} = \sigma_a^2 \frac{N_0}{\sum_{k=0}^{N_1} |M_k^+|^2}, \tag{34}$$

where the coefficients M_k^+ are determined from the following equation

$$\sum_{k=0}^{\infty} M_k^+ M_{m-k}^- = M_m \quad \text{for} \quad \text{all } m. \tag{35}$$

Here, $\{M_k^+\}$ is the causal "root" of the two sided sequence $\{M_k\}$. It satisfies $M_k^+ = 0$ for $k < 0$ and $M_k^+ = (M_{-k}^-)^*$ for $k \geq 0$. It is shown in Ref. 3 that

$$|M_0^+|^2 = N_0 \exp\left[\frac{T}{2\pi}\int_{-\pi/T}^{\pi/T} \ln\left(\frac{R(\omega)}{N_0} + 1\right) d\omega\right], \tag{36}$$

and when this formula is inserted into eq. (34) we get the well-known result for the decision feedback equalizer,

$$\epsilon_{\text{opt}} = \sigma_a^2 \exp\left[-\frac{T}{2\pi}\int_{-\pi/T}^{\pi/T} \ln\left(\frac{R(\omega)}{N_0} + 1\right) d\omega\right]. \tag{37}$$

Unfortunately, there is no similar simple expression for $|M_k^+|^2$, for $k \neq 0$; therefore, we are forced to numerically factor the two-sided sequence $\{M_k\}$ into its causal and anticausal root.

IV. DISCUSSION OF NUMERICAL RESULTS

In this section, the minimal mse of data-aided equalization is evaluated numerically for certain channels and for various sets of canceler taps. We will exhibit and discuss the behavior of tbe mse, $\epsilon_{\text{opt}}(N_1, N_2)$, as a function of N_1 and N_2 for typical telephone channels. As a point of reference, note the following easily proved inequalities:

$$\epsilon_{\text{opt}}(0, 0) \geq \epsilon_{\text{opt}}(0, \infty) \geq \epsilon_{\text{opt}}(\infty, \infty).$$

In the following, we examine three types of cancelers:

(*i*) Starting from the infinite length linear equalizer whose mse is $\epsilon_{opt}(0, 0)$, we increase the number of causal canceler taps; i.e., we examine $\epsilon_{opt}(0, N_2)$ for $N_2 = Q = 0, \cdots, 15$.

(*ii*) Starting from the infinite length linear equalizer, we increase the number of known data symbols alternating between causal and noncausal ones; i.e., we examine $\epsilon_{opt}(N_1, N_2)$ for $Q = N_1 + N_2 = 0, \cdots, 15$ and where $N_1 = N_2$ for Q even, $N_2 = N_1 + 1$ for Q odd.

(*iii*) Starting from the infinite length decision feedback equalizer whose mse is $\epsilon_{opt}(0, \infty)$, we examine the behavior when noncausal taps are added; i.e., $\epsilon_{opt}(N_1, \infty)$ for $N_1 = Q = 0, \cdots, 15$.

Equation (25) is used to determine the mse for cases (*i*) and (*ii*), where H_0 is obtained as the solution of eq. (26). To determine the sequence $\{M_k^{-1}\}$ for all k, we observe that the Fourier transform of the sequence $\{R_k\}$ is related to the overall transfer function of the channel as follows:

$$R(\omega) = \sum_{k=-\infty}^{\infty} \left| P\left(\omega - k\frac{2\pi}{T}\right)\right|^2. \tag{38}$$

Clearly, $R(\omega)$ is periodic with period $2\pi/T$, and it is only dependent on the magnitude of the overall channel transfer function, $P(\omega)$. Therefore, phase distortion in the channel has no effect on the mse. This is reflected in the well-known fact that phase distortion can be perfectly equalized without noise enhancement. Therefore, the sequence, $\{M_k^{-1}\}$, is obtained as follows

$$M_k^{-1} = \frac{T}{2\pi}\int_{-\pi/T}^{\pi/T} \frac{e^{jk\omega T}d\omega}{R(\omega) + N_0}, \tag{39}$$

where $j = \sqrt{-1}$. Fast Fourier transform techniques are used to evaluate eq. (39). Numerical tests show that it suffices to take 64–128 samples of $R(\omega) + N_0$ in the interval $[-\pi/T, \pi/T]$. The fact that the coefficient matrix in eq. (26) is positive definite and Toeplitz, makes it possible to obtain the desired solution, H_0, recursively. This is done with the Levinson algorithm.[11–14]

For case (*iii*), we evaluate eq. (34). The sequence, $\{M_k^{+}\}$, is obtained from the following approach.[3] First determine $\{F_k\}$ for all k such that

$$\ln(M(\omega)) = \sum_{k=-\infty}^{\infty} F_k e^{-jk\omega T}. \tag{40}$$

Then it follows that

$$M(\omega)^{+} = \exp\left\{\sum_{k=0}^{\infty} F_k e^{-jk\omega T}\right\}, \tag{41}$$

and

$$M_k^+ = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} M(\omega)^+ e^{jk\omega T} d\omega . \tag{42}$$

Fast Fourier transforms are used to obtain consecutively $\{F_k\}$, $M(\omega)$, and M_k^+. The overall channel power transfer function, $|P(\omega)|^2$, is assumed to consist of a raised cosine shaped transmitting filter with relative excess bandwidth, $\alpha = 0.15$,

$$|T(\omega)|^2 = \begin{cases} 1 & \text{for } |\omega| < (1-\alpha)\pi/T \\ 0.5\left[1 - \sin\left(\omega - \dfrac{\pi}{T}\right)\dfrac{T}{2\alpha}\right] & \text{for } (1-\alpha)\dfrac{\pi}{T} < \omega < (1+\alpha)\dfrac{\pi}{T} \\ 0.5\left[1 + \sin\left(\omega + \dfrac{\pi}{T}\right)\dfrac{T}{2\alpha}\right] & \text{for } -(1+\alpha)\dfrac{\pi}{T} < \omega < -(1-\alpha)\dfrac{\pi}{T} \\ 0 & \text{elsewhere} \end{cases}$$

and cascaded with the channel power transfer function $|G(\omega)|^2$.

Figure 3 shows two different channel power transfer functions, $|G(\omega)|^2$, which are used to derive the subsequent numerical results. In (a) we show the equivalent baseband transfer function for the worst channel meeting the basic conditions of private lines (BASICBAD).[15] Part (b) shows a transfer function (CABLE) with linearly increasing attenuation. The parameters P_1 and P_2 indicate the attenuation at $\omega = -\pi/T$ and $\omega = \pi/T$. A model for a baseband cable channel is obtained when $P_1 = P_2$.

Figure 4 shows the mse as a function of the number of canceler taps for the various channel transfer functions and for s/n of 20 dB at the receiver input. The dotted line represents type (*i*) canceler; the dashed line, type (*ii*); and the solid line, type (*iii*). The curves for types (*i*) and (*ii*) start at the minimal mse for the infinite length equalizer and the curve for type (*iii*) starts at the minimal mse of the infinite decision feedback equalizer.

As can be observed, all curves converge very rapidly to their asymptotes. The curve for type (*i*) indicates that only 3 causal coefficients suffice to closely approximate the performance of an infinite decision feedback equalizer. The curve for type (*ii*) suggests that a total of 6 coefficients (3 causal and 3 anticausal) results in a performance which is very close to the optimal (the matched-filter bound). The curve for type (*iii*) reaches very close to the mse obtained from the matched-filter bound with only 3 noncausal coefficients, in addition to an infinite decision feedback equalizer. These results are virtually independent of the channel involved.

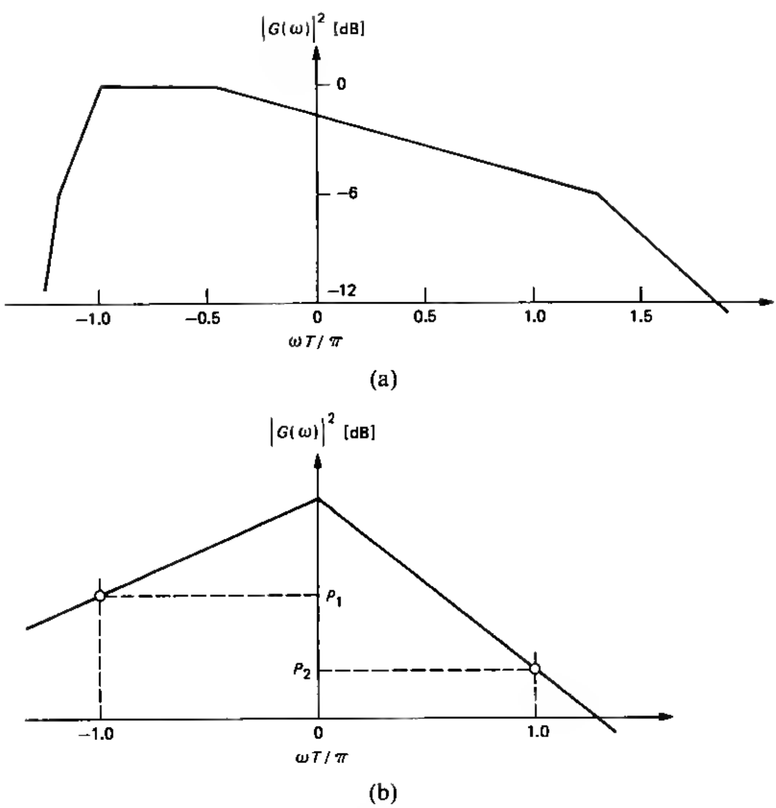

Fig. 3—Power transfer functions. (a) BASICBAD channel; (b) CABLE channel.

The channel, however, influences the best mse which is obtainable with the infinite equalizer, i.e. $\epsilon_{\text{opt}}(0, 0)$, and with the infinite decision feedback equalizer, i.e. $\epsilon_{\text{opt}}(0, \infty)$. Table I shows these figures for the various channels.

The minimal mse obtained from the matched-filter bound is −20.04 dB. Therefore, 1.8 to 4.7 dB can be gained for the channels considered if a canceler of three causal and three noncausal coefficients is included.

APPENDIX A

Solution of an Infinita Set of Equations with Finite Gaps

Let

$$M_k = R_k + N_0 \delta_{k,0}\,, \tag{43}$$

and consider

$$\sum_{k \notin J} U_k M_{m-k} = (1 - U_0) R_m \quad \text{for} \quad m \notin J\,, \tag{44}$$

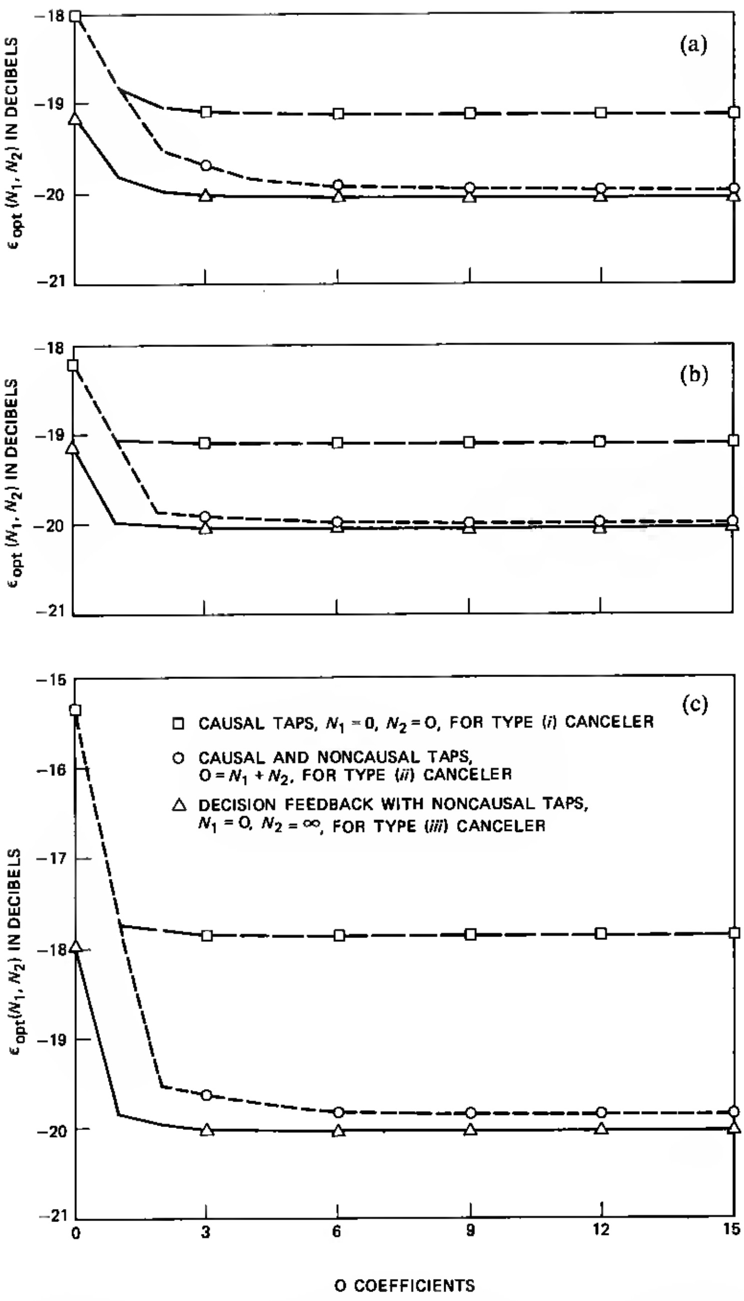

Fig. 4—Mean-square error for data-aided equalizations. (a) BASICBAD channel; (b) CABLE channel, $P_1 = P_2 = -10$ dB; (c) CABLE channel, $P_1 = -10$ dB, $P_2 = -20$ dB.

Table I—Minimal mean-square error

Mean-Square Error Channel	∞ Equalizer ϵ_{opt} (0, 0) in dB	∞ Decision Feedback Equalizer ϵ_{opt} (0, ∞) in dB
BASICBAD	−18.05	−19.16
CABLE		
$P_1 = 10, P_2 = 10$	−18.22	−19.13
$P_1 = 10, P_2 = 20$	−15.35	−17.97

where J is a finite set. It always contains the number zero but is otherwise arbitrary. Equation (44) is an infinite set of equations with a finite gap in both the indices of tbe unknowns U_k and the right-hand sides R_m. Notice that eq. (44) reduces to a discrete convolution and, therefore, is easy to solve if k and m are allowed to take on all the integers, or if the gap could be removed somehow.

Now consider instead of eq. (44) the following set of equations

$$\sum_{k=-\infty}^{\infty} V_k M_{m-k} = (1 - U_0)(R_m - H_m) \quad \text{for} \quad \text{all } m\,, \tag{45}$$

which is a discrete convolution. In order that eq. (45) conform to eq. (44), the auxiliary sequence $\{H_m\}$ must satisfy

$$H_m = 0 \quad \text{for} \quad m \notin J\,, \tag{46}$$

and

$$V_m = 0 \quad \text{for} \quad m \in J\,. \tag{47}$$

To accomplish this, the values of H_m for $m \in J$ are determined such that these constraints are forced to be satisfied. This is always possible since there are $N_1 + N_2 + 1$ free parameters, H_m, and the same number of conditions on V_m. From the above, it follows that

$$U_k = V_k \quad \text{for} \quad k \notin J\,. \tag{48}$$

This is easily proved by subtracting eq. (45) for $m \notin J$ from eq. (44). Now define the sequence $\{M_k^{-1}\}$ sucb that

$$\sum_{k=-\infty}^{\infty} M_k M_{m-k}^{-1} = \delta_{m,0}\,. \tag{49}$$

It can then be shown that

$$V_k = (1 - U_0) \sum_{m=-\infty}^{\infty} (R_m - H_m) M_{k-m}^{-1} \quad \text{for} \quad \text{all } k\,. \tag{50}$$

Since $V_k = 0$ for $k \in J$, we conclude from eq. (50) that

$$\sum_{m=-\infty}^{\infty} H_m M_{k-m}^{-1} = \sum_{m=-\infty}^{\infty} R_m M_{k-m}^{-1} \quad \text{for} \quad k \in J\,,$$

and since $H_m = 0$ for $m \notin J$, it follows that

$$\sum_{m \in J} H_m M_{k-m}^{-1} = \delta_{k,0} - N_0 M_k^{-1} \quad \text{for} \quad k \in J\,, \tag{51}$$

where we used eq. (43) and eq. (49) for the right-hand side. Eq. (51) can be solved if the Toeplitz matrix generated by the sequence $\{M_k^{-1}\}$ is not singular. This is always the case when $M(\omega) = R(\omega) + N_0$ is bounded away from zero and infinity, i.e. for all systems of practical interest.[10]

For evaluation of the mse, we need

$$\sum_{k \notin J} U_k R_{-k} = \sum_k V_k R_{-k}\,, \tag{52}$$

wbere tbe equation holds because $V_k = 0$ for $k \in J$. Since J always contains the number zero, we conclude from eq. (43) that R_{-k} can be replaced by M_{-k}. This yields togetber with eq. (52) and eq. (45)

$$\sum_{k \notin J} U_k R_{-k} = (1 - U_0)(R_0 - H_0)\,. \tag{53}$$

We use eq. (53) together with eq. (22) to obtain

$$U_0 = \frac{H_0}{N_0 + H_0}\,, \tag{54}$$

which finally leads to

$$\epsilon_{\text{opt}} = \sigma_a^2 \frac{N_0}{N_0 + H_0}\,, \tag{55}$$

the desired main result.

APPENDIX B

Analysis of the One-Sided Canceler of Infinite Length

For N_1 finite and $N_2 = \infty$, the set of equations (24) which determines U_k reads as follows:

$$\sum_{k<-N_1} U_k M_{m-k} = (1 - U_0) R_m \quad \text{for} \quad m < -N_1\,. \tag{56}$$

To solve this one-sided convolution, we factor the sequence $\{M_m\}$ into a causal part $\{M_m^+\}$ and an anticausal part $\{M_m^-\}$; i.e.,

$$M_m = \sum_{n=0}^{\infty} M_n^+ M_{m-n}^-\,, \tag{57}$$

where

$$M_n^+ = 0 \quad \text{for} \quad n < 0 \tag{58a}$$

$$M_n^- = 0 \quad \text{for} \quad n > 0\,. \tag{58b}$$

We now define a sequence Y_n such that

$$\sum_{m=0}^{\infty} M_n^+ Y_{m-n} = (1 - U_0) R_m \quad \text{for all} \quad m\,. \tag{59}$$

Now insert eq. (57) into eq. (56) to obtain

$$\sum_{k<-N_1} U_k \sum_{n=0}^{\infty} M_n^+ M_{m-k-n}^- = (1 - U_0) R_m \quad \text{for} \quad m < -N_1\,. \tag{60}$$

In addition, substract eq. (60) from eq. (59) for $m < -N_1$ and obtain the following set of equations

$$Y_m = \sum_{k<-N_1} U_k M_{m-k} \quad \text{for} \quad m < -N_1\,. \tag{61}$$

Multiply eq. (61) by M_m^+, sum over all $m < -N_1$, and use eq. (57) on the right-hand side to obtain

$$\sum_{k<-N_1} Y_m M_{-m}^+ = \sum_{k<-N_1} U_k M_{-k}\,. \tag{62}$$

Recall that

$$M_k = R_k + N_0 \delta_{k,0}\,, \tag{63}$$

and compare eqs. (57) and (59) which determines Y_m as

$$Y_m = (1 - U_0) M_m^- \quad \text{for} \quad m \neq 0\,. \tag{64}$$

Now insert eq. (64) into eq. (62) and make use of eq. (63) once more to obtain the one-sided sum which is required in eq. (53):

$$\sum_{k<-N_1} R_{-k} U_k = (1 - U_0) \sum_{m<-N_1} M_m^- M_{-m}^+\,. \tag{65}$$

Since $M_m^{-*} = M_{-m}^+$, it can be shown that

$$\begin{aligned}\sum_{k<-N_1} R_{-k} U_k &= (1 - U_0) \sum_{m<-N_1} |M_m^-|^2 \\ &= (1 - U_0) \sum_{m>N_1} |M_m^+|^2.\end{aligned} \tag{66}$$

Also, insert eq. (66) into eq. (53) and use eq. (63) to find

$$N_0 + H_0 = M_0 - \sum_{m>N_1} |M_m^+|^2. \tag{67}$$

With eq. (57) evaluated for $m = 0$, we finally obtain

$$N_0 + H_0 = \sum_{m=0}^{N_1} |M_m^+|^2, \tag{68}$$

and with eq. (55) we get our desired result,

$$\epsilon_{\text{opt}} = \sigma_a^2 \frac{N_0}{\sum_{m=0}^{N_1} |M_m^+|^2}. \tag{69}$$

REFERENCES

1. R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communication,* New York: McGraw-Hill, 1968.
2. T. Berger and D. W. Tufts, "Optimal Pulse Amplitude Modulation, Part I. Transmitter-Receiver Design and Bounds from Information Theory," IEEE Trans. Information Theory, *IT-13,* No. 2 (April 1967), pp. 196–208.
3. J. Salz, "Optimum Mean-Square Decision Feedback Equalization," B.S.T.J., *52,* No. 8 (October 1973), pp. 1341–73.
4. D. D. Falconer and G. J. Foschini, "Theory of Minimum Mean-Square-Error QAM Systems Employing Decision Feedback Equalization," B.S.T.J., *52,* No. 10 (December 1973), pp. 1821–49.
5. T. M. Cover, J. F. Hayes, and J. Riera, unpublished work.
6. T. Kailath, "A General Likelihood-Ratio Formula for Random Signals in Gaussian Noise," IEEE Trans. Information Theory, *IT-15,* No. 3 (May 1969), pp. 350–61.
7. G. D. Forney, Jr., "Maximum-Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference," IEEE Trans. Information Theory, *IT-18* (May 1972), pp. 363–78.
8. D. D. Falconer and F. R. Magee, Jr., "Adaptive Channel Memory Truncation for Maximum Likelihood Sequence Estimation," B.S.T.J., *52,* No. 9 (November 1973), pp. 1541–62.
9. D. D. Falconer, "Jointly Adaptive Equalization and Carrier Recovery in Two-Dimensional Digital Communication Systems," B.S.T.J., *55,* No. 3 (March 1976), pp. 317–34.
10. U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications,* Berkeley: University of California Press, 1958.
11. B. D. O. Anderson and J. B. Moore, *Optimal Filtering,* Englewood Cliffs: Prentice-Hall, 1979.
12. N. Levinson, "The Wiener RMS (Root Mean Square) Error Criterion in Filter Design and Prediction," J. Math. and Phys., *25* (1946), pp. 261–78.
13. W. F. Trench, "Weighting Coefficients for the Prediction of Stationary Time Series from the Finite Past," SIAM J. Appl. Math., *15,* No. 6 (1967), pp. 1502–10.
14. S. Zohar, "The Solution of a Toeplitz Set of Linear Equations," J. Assoc. Computing Machinery, *21,* No. 2 (April 1974), pp. 272–6.
15. Bell System Technical Reference, "Data Communications Using Voiceband Private Line Channels," October 1973, Table 3, p. 13.